Chapter 5: Torus Networks

Outline:

- Torus networks
  - Torus and butterfly comparison
- Performance
  - Throughput
  - Latency
  - Path diversity
- Building mesh and torus networks
Torus Networks

➢ Torus and mesh: (fig. 5.1 and 5.2)
  - k-ary n-cube (torus), k-ary n-mesh (mesh)
  - \( N = k^n \) nodes
  - regular n-dimensional grid
  - k nodes in each dimension
  - dimensions may have different radix (fig. 5.3)
    • (Radix: the number of inputs and outputs of each switching node)
Torus and butterfly comparison

Table 1: Comparison between torus and butterfly topologies

<table>
<thead>
<tr>
<th>Torus</th>
<th>Butterfly</th>
</tr>
</thead>
<tbody>
<tr>
<td>k-ary n-cube</td>
<td>k-ary n-fly</td>
</tr>
<tr>
<td>Short wires, high speed without repeaters</td>
<td>Long wires, repeaters might be needed</td>
</tr>
<tr>
<td>Good path diversity</td>
<td>No path diversity</td>
</tr>
<tr>
<td>Good load balance</td>
<td>Load imbalance may occur</td>
</tr>
<tr>
<td>Bidirectional (can be also unidirectional)</td>
<td>Unidirectional</td>
</tr>
<tr>
<td>Direct</td>
<td>Indirect / isomorphic</td>
</tr>
<tr>
<td>Large hop count</td>
<td>Same hop count for any packet</td>
</tr>
<tr>
<td>Many routing possibilities</td>
<td>Simple routing</td>
</tr>
<tr>
<td>Edge symmetric</td>
<td>Channels between nodes are symmetric</td>
</tr>
<tr>
<td>Physical locality</td>
<td>Unable to exploit such locality</td>
</tr>
</tbody>
</table>
Performance

Characterized by its:

- throughput
- latency
- path diversity
Throughput

- Limited by pin bandwidth or bisection bandwidth
  - (bisection: the cut that divides the entire network nearly in a half)
- Torus dimension has to do with the best throughput:
  - dimension must be high enough to keep the bisection limited
  - dimension must be small enough to keep all wires below the critical wire length
- Results in Table 5.1 show that
  - throughput increases with dimension before the network becomes bisection limited (at n = 3)
  - n >= 3, throughput is the same since the design uses all the available bisection wiring
Latency

- Depends on dimension:
  - at low dimension, high hop count dominates the latency
  - at high dimensions, serialization latency dominates due to the narrow channel width
- A low, intermediate dimension usually results in optimal latency
- As Table 5.2 illustrates, the latency is smallest when the dimension is 4
  - considering also throughput, dimension 3 could be considered as well
Path diversity

- Path diversity: Distinct paths between the source and terminal nodes
  - load balancing
  - robustness, networks can be reconfigured around faulty channels
- Increases rapidly with dimension
- You can calculate only minimal paths or all paths
Building mesh and torus networks

Network nodes have to be mapped to their place in the physical space:
- 1D, 2D, or 3D depending on the packaging technology

Tori and meshes are relatively easy to map
- short wires
- mesh is the simplest case when its dimension equals to the physical dimension of the packaging technology

Folding (Fig. 5.5 and 5.6)
- if there is long end-around channels
  - doubles the length of other channels